

NanoTrack – Sistema Inteligente de Gerenciamento de Dados de Síntese de Nanoestruturas

André S. F. Silva¹, Paulo H. S. Batista¹, Thamara K. C. Andrade¹,
Omar P. Vilela Neto¹, Sérgio Campos¹, Alessandra Faria-Campos¹

¹Departamento de Ciência da Computação – Universidade Federal de Minas Gerais
(UFMG) Belo Horizonte – MG – Brazil

{andresfs, paulo.batista, omar, scampos, alessa, thamara}@dcc.ufmg.br

Abstract. *This work describes a proposal for a tool to support nanoscience and nanotechnology laboratories based on the concept of LIMS (Laboratory Information Management Systems). The tool consists of a workflow capable of storing and managing data related to the synthesis of nanostructures, and resources for data sharing and intelligent systems to infer parameters to be used in future synthesis of new nanostructures.*

Resumo. *Este trabalho descreve uma proposta de ferramenta para apoio a laboratórios de nanociência e nanotecnologia baseada no conceito de LIMS (Laboratory Information Management Systems). A ferramenta consiste de um workflow capaz de armazenar e gerenciar os dados relacionados à síntese de nanoestruturas, além de recursos para o compartilhamento de dados e sistemas inteligentes para inferir parâmetros a serem utilizados em sínteses de futuras nanoestruturas.*

1. Introdução

O interesse em nanociência e nanotecnologia tem crescido significativamente devido à possibilidade de desenvolvimento de novos materiais com propriedades superiores e inovadoras. A nanotecnologia, em um futuro breve, poderá mudar a natureza de vários objetos produzidos pelos seres humanos, gerando um impacto nas mais diversas áreas, tais como medicina, engenharia, telecomunicações, energia, computação, dentre outras [Dias et al. 2011].

A pesquisa em nanotecnologia, contudo carece ainda de ferramentas computacionais para apoiá-la, como é mostrado em [de la Iglesia et al. 2011]. As atividades envolvidas na síntese de nanoestruturas são em geral complexas, o tempo despendido na experimentação é grande e as informações associadas aos experimentos são armazenadas de forma precária.

A obtenção de dados para a fabricação de qualquer composto nanométrico é dificultada por fatores tecnológicos e econômicos e esta falta de dados precisos sobre os experimentos com nanomateriais restringe a descoberta de conhecimento e dificulta a criação de nanomateriais inovadores [de la Iglesia et al. 2011]. Uma solução para este problema é a utilização de sistemas computacionais para a gerência destes dados. Os LIMS (Laboratory Information Management Systems) são um dos tipos de sistemas que podem ser usados para esta finalidade.

Os LIMS são sistemas desenvolvidos com o objetivo de facilitar o gerenciamento e armazenamento de dados de laboratório, contemplando todo o ciclo de vida dos dados (coleta de dados, armazenamento, análise, geração de relatórios e arquivamento) e permitindo o acesso dos usuários aos respectivos resultados armazenados [Hinton 1995].

Devido à complexidade dos processos gerenciados, os LIMS costumam ser específicos para cada laboratório, sendo difícil o uso de um mesmo LIMS por laboratórios diferentes devido as diferenças nos processos de cada um. Assim, a integração de laboratórios diferentes com um mesmo sistema de gerenciamento é um dos desafios da utilização destes. Portanto, é essencial o desenvolvimento de LIMS customizados para a gerência de processos científicos. Entre estes processos está a fabricação de compostos nanométricos. O objetivo principal deste trabalho é construir um sistema para gerenciamento de dados de experimentos em nanotecnologia, com uma plataforma de apoio ao pesquisador que possua funcionalidades adequadas à captura e análise de dados desta natureza, como atividades de síntese, ambiente de simulação e compartilhamento de informação entre pesquisadores.

2. Trabalhos relacionados

Vários algoritmos e ferramentas computacionais para gerenciamento de dados de laboratório têm sido desenvolvidos, mas sistemas destinados ao uso de laboratórios de nanotecnologia são raros e pouco flexíveis [de la Iglesia et al. 2011]. Diante do avanço constante e rápido da área, projetar um software que seja capaz de resistir a estes avanços constitui um desafio. Dois trabalhos apresentados na literatura [Wermelinger et al. 2008] e [Crnkovic et al. 2002] exploraram os princípios dos projetos conhecidos por impactar na manutenção de software e planejaram um modelo de modularização de ferramenta baseado nestes princípios para permitir a evolução da ferramenta através de plugins simples e ajustáveis.

Neste trabalho, propomos a construção de um workflow para o gerenciamento e análise de dados de experimentos de nanotecnologia utilizando o sistema Flux®, uma extensão do sistema SIGLa [Melo et al. 2010], o qual utiliza uma abordagem baseada em workflows. Um workflow pode ser definido como a sequência de passos e tarefas executadas sequencialmente, diante de um conjunto de regras e procedimentos que devem ser respeitados para se concluir um processo [Hollingsworth et al. 1995].

Para auxiliar os especialistas na síntese de futuras nanoestruturas, este trabalho propõe uma ferramenta com sistemas inteligentes de inferência e previsão integrados ao workflow como plugins que permitirão a aplicação de técnicas específicas de inferência quando necessário. A abordagem orientada a workflows do NanoTrack permite que o processo e a aplicação evoluam paralelamente e de maneira independente, pois a lógica do processo está totalmente separada da lógica da aplicação. Isto também provê uma grande flexibilidade, pois, caso o processo de fabricação mude ou haja a necessidade de se armazenar qualquer outro dado a única coisa que precisa ser alterada é o workflow, e não o código fonte. Além disso, essa abordagem irá possibilitar que a ferramenta acompanhe a evolução das técnicas e facilitará a introdução de novas técnicas.

Para a realização dos experimentos e validação do algoritmo serão utilizados a mesma amostra de dados do trabalho de [Singulani et al. 2008] e dados coletados junto aos Institutos Nacionais de Ciência e Tecnologia de Nanomateriais de Carbono e Nano-

dispositivos Semicondutores utilizados na síntese de nanodispositivos. O objetivo da utilização de diferentes conjuntos de amostras é demonstrar que o sistema desenvolvido não é especializado na resolução de um único problema, cobrindo diversos projetos de síntese de nanoestruturas.

3. Metodologia

A modelagem do workflow usado no NanoTrack está sendo feita através da aplicação Together Workflow Editor, um editor de workflows gráfico gratuito que gera um arquivo XML que representa o workflow. Cada etapa do processo (Síntese, Microscopia AFM, Difração de Raio X, etc) é representada dentro do workflow como uma atividade e será exibida pelo NanoTrack como uma etapa do processo.

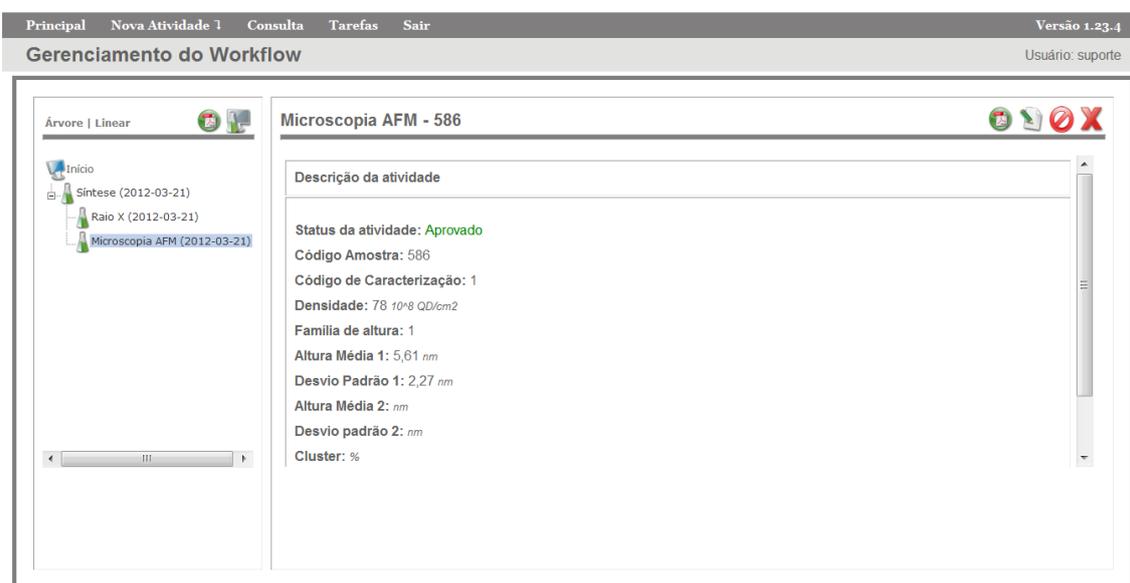


Figure 1. Interface do ambiente NanoTrack.

Para cada dado acerca daquela atividade existe um atributo correspondente dentro do sistema. A adição de novas atividades ao processo e novos atributos para novas atividades ou atividades existentes é feita de forma simplificada no editor.

Após a etapa de modelagem do processo, o workflow é carregado no NanoTrack, que oferece uma interface gráfica amigável ao usuário que pode ser acessada através dos navegadores web, vide Figura 1. O acesso ao sistema via web, permite a entrada de dados a partir de qualquer dispositivo que possua acesso à rede e um navegador web compatível. Através da utilização de um ambiente integrado pelos pesquisadores, uma base de dados de diversos projetos será construída na medida que eles forem armazenados.

Estes dados serão utilizados na realização de sistemas de inferências para prever novos experimentos laboratoriais, visando à síntese de nanodispositivos aprimorados. Os algoritmos de inferência podem ser implementados de diferentes maneiras e serão integrados ao workflow de forma dinâmica, através de plugins. Contudo, dependendo da nanoestrutura a ser prevista, um sistema de inferência pode ser mais eficiente do que outro. Em alguns casos, uma combinação de sistemas de inferência pode ser ainda mais eficiente. Desta forma, este trabalho implementará um comitê que será responsável por

receber os resultados de todos os sistemas de inferência selecionados e disponibilizar os melhores resultados ao usuário. Na medida em que o workflow for recebendo novas amostras, é esperado que os sistemas de inferência e o comitê sejam mais precisos.

Inicialmente, pelo menos dois algoritmos de inferência irão compor a ferramenta proposta, baseados nos trabalhos [Dias et al. 2011] e [Singulani et al. 2008]. A ideia é demonstrar flexibilidade na inferência de resultados para os problemas, demonstrando que as soluções podem ser conjuntas dentre os algoritmos disponíveis.

4. Resultados esperados

O uso de um sistema computacional para gerenciar os dados de fabricação de nanoestruturas irá prover uma plataforma de armazenamento com permissão de acesso aos dados para os usuários autorizados e protegendo os mesmos de acessos não autorizados. Com a construção deste sistema, esperamos possibilitar uma melhora na qualidade e velocidade em que os dados são armazenados. O uso de uma ferramenta automática possibilitará isto, uma vez que qualquer dado inserido passará por um processo de validação automático e a análise dos dados inseridos poderá também ser feita de forma automática através dos plugins desenvolvidos. Esperamos com o desenvolvimento do NanoTrack suprir as principais necessidades em gerenciamento de dados de laboratórios de nanotecnologia, provendo uma ferramenta eficiente e segura para a utilização por pesquisadores da área.

References

- Crnkovic, I., Larsson, S., and Stafford, J. (2002). Component-based software engineering: building systems from components at 9th iee conference and workshops on engineering of computer-based systems. *ACM SIGSOFT Software Engineering Notes*, 27(3):47–50.
- de la Iglesia, D., Harper, S., Hoover, M., Klaessig, F., Lippell, P., Maddux, B., Morse, J., Nel, A., Rajan, K., Reznik-Zellen, R., et al. (2011). Nanoinformatics 2020 roadmap.
- Dias, D. M., Singulani, A. P., PACHECO, M. A. C., SOUZA, P. L., Pires, M., and Vilela Neto, O. P. (2011). Self-assembly quantum dots growth prediction by quantum-inspired linear genetic programming. In *Evolutionary Computation (CEC), 2011 IEEE Congress on*, pages 2075–2082. IEEE.
- Hinton, M. (1995). Laboratory management systems. *New York*.
- Hollingsworth, D. et al. (1995). Workflow management coalition: The workflow reference model. *Document Number TC00-1003*, (1.1).
- Melo, A., Faria-Campos, A., DeLaat, D. M., Keller, R., Abreu, V., and Campos, S. (2010). Sigla: an adaptable lims for multiple laboratories. *BMC genomics*, 11(Suppl 5):S8.
- Singulani, A., Vilela Neto, O., Aurélio Pacheco, M., Vellasco, M., Pires, M., and Souza, P. (2008). Computational intelligence applied to the growth of quantum dots. *Journal of Crystal Growth*, 310(23):5063–5065.
- Wermelinger, M., Yu, Y., and Lozano, A. (2008). Design principles in architectural evolution: a case study. In *Software Maintenance, 2008. ICSM 2008. IEEE International Conference on*, pages 396–405. IEEE.